Convex Sets. Convex functions. D.C. functions

Liudmila Poliakova

Saint-Peterburg State University,

Department of Applied Mathematics and Control Processes,
Saint-Petersburg, Russia

2018

Convex sets

A set $X \subset \mathbb{R}^n$ is called convex, if for all $x_1 \in X$ and $x_2 \in X$ the next formula

$$\lambda x_1 + (1-\lambda)x_2 \in X \quad \forall \lambda \in [0,1] \subset \mathbb{R},$$

hold.

We assume that the empty set $\emptyset$ is convex by definition.

The sum of two convex sets $X_1, X_2 \subset \mathbb{R}^n$ is called the set

$$X = X_1 + X_2 = \{x_1 + x_2 \mid x_1 \in X_1, \quad x_2 \in X_2\}.$$

The set $X = X_1 + X_2$ is called the algebraic sum of two convex sets X_1 and X_2 or Minkowski's sum.

By writing $X_1 - X_2$ we will understand the set $X_1 + (-X_2)$.

Convex sets

A set $X \subset \mathbb{R}^n$ is called convex, if for all $x_1 \in X$ and $x_2 \in X$ the next formula

$$\lambda x_1 + (1-\lambda)x_2 \in X \quad \forall \lambda \in [0,1] \subset \mathbb{R},$$

hold.

We assume that the empty set $\emptyset$ is convex by definition.

The sum of two convex sets $X_1, X_2 \subset \mathbb{R}^n$ is called the set

$$X = X_1 + X_2 = \{x_1 + x_2 \mid x_1 \in X_1, \quad x_2 \in X_2\}.$$

The set $X = X_1 + X_2$ is called the algebraic sum of two convex sets X_1 and X_2 or Minkowski's sum.

By writing $X_1 - X_2$ we will understand the set $X_1 + (-X_2)$.

If the set $X \subset \mathbb{R}^n$ is convex, then its every scalar multiple $\alpha X, \alpha \in \mathbb{R}$,

1. Let I be an arbitrary set of indices i, a set $X_i \subset \mathbb{R}^n$ is convex for each index $i \in I$. Then

$$X = \bigcap_{i \in I} X_i$$

is convex as well.

2. Let sets $X_1 \subset \mathbb{R}^n$ and $X_2 \subset \mathbb{R}^n$ be convex, then the algebraic sum $X_1 + X_2$ is also convex.

Let $f : \mathbb{R}^n \to \mathbb{R} \bigcup \{+\infty\} \bigcup \{-\infty\}$.
The set

$$\text{dom } f = \{x \in \mathbb{R}^n \mid \ f(x) < +\infty\}$$

is called the effective domain of a convex function f.
The set

$$\text{epi } f = \{[x, \mu] \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leqslant \mu\}$$

is called the epigraph of f.

A convex function f is said to be proper if its epigraph is non-empty and contains no vertical lines, i.e. if $f(x) < +\infty$ for at least one x and $f(x) > -\infty$ for every x.
For proper convex functions it is possible to give another definition which equivalent to the above.

A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is called convex if the next relation

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leqslant \lambda_1 f(x_1) + \lambda_2 f(x_2) \quad \forall x_1, x_2 \in \mathbb{R}^n,$$

$$\lambda_1, \lambda_2 \geq 0, \; \lambda_1 + \lambda_2 = 1.$$

holds.

Let f be a convex function.

If $x \in \text{int dom} f$, then f is continuous at x.

Let f be a convex function. If the partial derivatives of f with respect to each variable exist at a point $x \in \text{int dom} f$ then f is differentiable at x.

The conjugate of a function f is

$$f^*(v) = \sup_{x \in \mathbb{R}^n} \{\langle x, v\rangle - f(x)\}, \ v \in \mathbb{R}^n.$$

Obviously, that the equality

$$f^*(v) = \sup_{x \in \text{dom } f} \{\langle x, v\rangle - f(x)\}, \ v \in \mathbb{R}^n$$

is true.

Note some of their properties.

1. f^* is closed and convex (even when f is not).

2. $f(x) + f^*(v) \geq \langle x, v \rangle \; \forall x \in \mathbb{R}^n, \forall v \in \mathbb{R}^n$

3. If f is a closed proper convex function then f^* is also a closed proper convex function and the equality

$$f(x) = f^{**}(x)$$

is true.

Examples of conjugate functions

1. Let

$$f(x) = \langle a, x\rangle + b, \quad a, x \in \mathbb{R}^n, \quad b \in \mathbb{R},$$

then

$$f^*(v) = \sup_{x\in\mathbb{R}^n} \{\langle x, v\rangle - \langle a, x\rangle - b\} = \sup_{x\in\mathbb{R}^n} \{\langle x, v - a\rangle - b\} =$$

$$= \sup_{x\in\mathbb{R}^n} \{\langle x, v - a\rangle\} - b.$$

Therefore

$$f^*(v) = \begin{cases} -b, & v = a, \\ +\infty, & v \neq a. \end{cases}$$

2. Let $f(x) = |x|,\ x \in \mathbb{R}$. then

$$f^*(v) = \sup_{x\in\mathbb{R}}\{x, v - |x|\} = \sup\{\sup_{x\geq 0}\{x, v - x\}, \sup_{x<0}\{x, v + x\}\}.$$

$$\sup_{x\geq 0}\{x, v - x\} = \begin{cases} 0, & v \leq 1, \\ +\infty, & v > 1. \end{cases}$$

$$\sup_{x<0}\{x, v + x\} = \begin{cases} 0, & v \geq -1, \\ +\infty, & v < -1. \end{cases}$$

Therefore,

$$f^*(v) = \begin{cases} 0, & |v| \leq 1, \\ +\infty, & |v| > 1. \end{cases}$$

The function conjugate to the function $f(x) = |x|$ is the indicator function of the segment $[-1, 1] \subset \mathbb{R}$.

3. Let

$$f(x) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c, \quad x, b \in \mathbb{R}^n,\ c \in \mathbb{R},$$

where A is a positive definite matrix. Then

$$f^*(v) = \frac{1}{2}\langle A^{-1}(v - b), v - b\rangle - c, \quad v \in \mathbb{R}^n.$$

A proper convex closed function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is convex if and only if when

$$f(x) = (f^*)^*(x), \quad x \in \mathbb{R}^n.$$

In this case $\mathrm{dom} f^* \neq \emptyset$.

Let f be a real-valued function defined on $\mathbb{R}^n$. Then f is said to have a directional derivative $f'(x,g)$ at $x \in \text{dom } f$ in the direction $g \in \mathbb{R}^n$ if

$$f'(x,g) = \lim_{\alpha \downarrow 0} \frac{f(x+\alpha g) - f(x)}{\alpha}$$

Theorem

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a convex function $x \in \text{dom } f$. If x is an interior point, then $f'(x,g) < +\infty$ exists for every $g \in \mathbb{R}^n$ and

A convex function is not necessarily differentiable.
Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be convex and $x \in \text{dom } f$.
The set

$$\partial f(x) = \{v \in \mathbb{R}^n \mid f(z) - f(x) \geq \langle v, z - x \rangle \quad \forall z \in \mathbb{R}^n\}$$

is called a subdifferential of f at x.
A vector $v \in \partial f(x)$ is called a subgradient of f at x.
A function f is called subdifferentiable if it is subdifferentiable at all $x \in \text{dom} f$.

If f is convex and differentiable, then its gradient at x is a subgradient.

Thus, the concept of subdifferential is defined only for points of the effective area, i.e., if $x \notin \text{dom } f$, then $\partial f(x) = \emptyset$.

Here is an example of a function that is not subdifferentiable at the boundary point of the effective area.
Consider the function $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$

$$f(x) = \begin{cases} -\sqrt{1-x^2}, & |x| \le 1, \\ +\infty, & |x| > 1. \end{cases}$$

At the point $x_0 = 1$ the set $\partial f(x_0)$ is empty.

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a proper convex function, $x_0 \in \text{dom } f$, $\partial f(x_0) \neq \emptyset$. Then $\partial f(x_0)$ is convex and closed.

For convex functions there is a close connection between the subdifferential and the derivative in the directions.

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function, $x_0 \in \text{dom } f$. Then

$$f'(x_0, g) = \sup_{v \in \partial f(x)} \langle v, g \rangle \quad \forall g \in \mathbb{R}^n \tag{1}$$

holds.

Let f_1, f_2 are proper convex functions, a point $x \in \text{int}\,(\text{dom}\ f_1) \cap \text{int}\,(\text{dom}\ f_2)$. Then for the function $f = f_1 + f_2$ the equality

$$\partial f(x) = \partial f_1(x) + \partial f_2(x) \tag{2}$$

is true.

Let $f(x) = \lambda f_1(x)$, where $\lambda > 0,\ x \in \mathbb{R}^n,\ f_1(x)$ is a proper convex function, then

$$\partial f(x) = \lambda \partial f_1(x).$$

Let f_1, f_2 be proper convex functions, a point $x \in \text{int}\,(\text{dom}\, f_1) \cap \text{int}\,(\text{dom}\, f_2)$. then for a function $f = \max\{f_1, f_2\}$ the equality

$$\partial f(x) = \text{co}\left\{ \bigcup_{i \in R(x)} \partial f_i(x) \right\} \tag{3}$$

is true, where $R(x) = \{i = 1, 2 \mid f_i(x) = f(x)\}$.

The set

$$\partial f(x) = \text{co} \left\{ \bigcup_{i \in R(x)} \partial f_i(x) \right\},$$

where $R(x) = \{i \in I \mid f_i(x) = f(x)\}$ is the subdifferential of the function

$$f(x) = \max_{i \in I} f_i(x), \quad I = 1, ..., m,$$

If a convex function f is differentiable at a point $x \in \mathbb{R}^n$, then

$$\partial f(x) = \{f'(x)\}.$$

Let f be a finite convex function on $\mathbb{R}^n$. Then the multivalued mapping

$$\partial f : \mathbb{R}^n \to 2^{\mathbb{R}^n}$$

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a proper convex function, $x \in \operatorname{dom} f$, $\partial f(x_0) \neq \emptyset$, $\varepsilon > 0$.
The set

$$\partial f(x) = \{v \in \mathbb{R}^n \mid \; f(z) - f(x) \geq \langle v, z - x \rangle \quad \forall z \in \mathbb{R}^n\}$$

is called ε-subdifferential of f at x.

Let f be a finite convex function on $\mathbb{R}^n$. Then the multivalued mapping

$$\partial_\varepsilon f : \mathbb{R}^n \times (0, +\infty) \to 2^{\mathbb{R}^n}$$

is continuous in the Hausdorff metric.

A point $x^* \in \mathbb{R}^n$ is a minimizer of a convex function f if and only if f is subdifferentiable at x and

$$0 \in \partial f(x^*)$$

Any local minimum of convex function is also a global minimum.

Let $f_1, f_2 : \mathbb{R}^n \to \mathbb{R}$ be finite convex functions on $\mathbb{R}^n$ and

$$f(x) = f_1(x) - f_2(x).$$

The function f is a quasidifferentiable function.

Quasidifferentiable function

Let a function f be defined on $\mathbb{R}^n$ and be directionally differentiable at a point $x \in \mathbb{R}^n$ and its directional derivative $f'(x,g)$ can be represented in the form

$$f'(x,g) = \lim_{\lambda \downarrow 0} \frac{f(x+\lambda g) - f(x)}{\lambda} = \max_{v \in \underline{\partial} f(x)} \langle v, g \rangle + \min_{w \in \overline{\partial} f(x)} \langle w, g \rangle,$$

where $\underline{\partial} f(x) \subset \mathbb{R}^n$, $\overline{\partial} f(x) \subset \mathbb{R}^n$ are convex compact sets in $\mathbb{R}^n$.

The function f is called a quasidifferential at a point $x \in \mathbb{R}^n$. A pair of sets

$$\mathcal{D}f(x) = [\underline{\partial} f(x), \overline{\partial} f(x)]$$

is called a quasidifferentiable of a quasidifferentiable function f at x. The set $\underline{\partial} f(x) \subset \mathbb{R}^n$ is called a subdifferential of f at x, the set $\overline{\partial} f(x) \subset \mathbb{R}^n$ is called a superdifferential of f at x. Differentiable, convex, concave functions, the maximum functions are quasidifferentiable functions.

As the function f is quasidifferentiable on $\mathbb{R}^n$ and

$$\mathcal{D}f(x) = [\partial f_1(x), -\partial f_2(x)]$$

is its quasidifferential at a point $x \in \mathbb{R}^n$, where $\partial f_i(x)$ are the subdifferentials of convex functions $f_i(x)$, $i = 1, 2$, at the point $x \in \mathbb{R}^n$ in the sense of convex analysis.

Let's consider the optimization problem: find

$$\inf_{x \in \mathbb{R}^n} f(x).$$

The following necessary optimality conditions for the function f on $\mathbb{R}^n$ hold.

Theorem.
For a point $x^* \in \mathbb{R}^n$ to be a minimizer of the function f on $\mathbb{R}^n$, it is necessary, that

$$\partial f_2(x^*) \subset \partial f_1(x^*). \tag{4}$$

For a point $x^* \in \mathbb{R}^n$ to be a maximizer of the function f on $\mathbb{R}^n$, it is necessary, that

$$\partial f_1(x^*) \subset \partial f_2(x^*). \tag{5}$$

If the inclusion

$$\partial f_2(x^*) \subset \text{int } \partial f_1(x^*) \tag{6}$$

holds at the point $x^* \in \mathbb{R}^n$ then this point is a strict local minimizer of the function f on $\mathbb{R}^n$.

If the inclusion

$$\partial f_1(x^*) \subset \text{int } \partial f_2(x^*)$$

is satisfied at the point $x^* \in \mathbb{R}^n$ then this point is a strict local maximizer of the function f on $\mathbb{R}^n$.

A point x^* is called a *inf*$-$ stationary point of f if inclusion (4) holds. A point x^{**} is called a *sup*$-$ stationary point of f if inclusion (5) holds.

J.-B. Hiriart-Urruty was the first who used ε-subdifferentials to obtain necessary and sufficient conditions of the global minimum of the difference of convex functions.

Theorem. *For a point $x^* \in \mathbb{R}^n$ to be a global minimizer of the function f on $\mathbb{R}^n$, it is necessary and sufficient that the inclusion*

$$\partial_\varepsilon f_2(x^*) \subset \partial_\varepsilon f_1(x^*) \quad \forall \varepsilon \geq 0 \tag{7}$$

be valid, where $\partial_\varepsilon f_i(x^*)$ are ε- subdifferentials of the convex functions f_i, $i = 1, 2$ at the point x^*.

We say that a point x^* is Clark's stationary point of function f on $\mathbb{R}^n$, if the next condition

$$\partial f_1(x^*) \bigcap \partial f_2(x^*) \neq \emptyset$$

holds. It is obvious that inf $-$ and sup-stationary points of the function f on $\mathbb{R}^n$ are also Clark's stationary points of f on $\mathbb{R}^n$.

Denote by

$$f^o(v) = f_2^*(v) - f_1^*(v), \quad v \in \mathbb{R}^n.$$

If a point $v \notin \text{dom } f_1^* \cup \text{dom } f_2^*$, then we face with the case $+\infty - \infty$.

Therefore in different cases under considering of certain extremal properties, we will define this function on different depending on the situation

1. If the point $x^* \in \mathbb{R}^n$ is a Clark's stationary point of the function f on $\mathbb{R}^n$, then the equality

$$f(x^*) = f^o(v) \quad \forall v \in \partial f_1(x^*) \cap \partial f_2(x^*)$$

holds.

2. If the point $x^* \in \mathbb{R}^n$ is an inf-stationary point of the function f on $\mathbb{R}^n$, then the equality

$$f(x^*) = f^o(v^*) \quad \forall v^* \in \partial f_2(x^*) \tag{8}$$

holds.

3. If the point $x^* \in \mathbb{R}^n$ is a sup-stationary point of the function f on $\mathbb{R}^n$, then the equality

$$f(x^*) = f^o(v^*) \quad \forall v^* \in \partial f_1(x^*) \tag{9}$$

holds.

Example 1.

Consider a function

$$f(x) = f_1(x) - f_2(x) = |x_1| - |x_2|, \quad x = (x_1, x_2) \in \mathbb{R}^2.$$

However, it is Clark's stationary point of the function f on $\mathbb{R}^2$.
Define the conjugate functions of f_1 and f_2.
We have

$$f_1^*(v) = \begin{cases} 0, & v \in \text{co}\ \{(1,0),(-1,0)\}, \\ +\infty, & v \notin \text{co}\ \{(1,0),(-1,0)\}, \end{cases}$$

$$f_2^*(v) = \begin{cases} 0, & v \in \text{co}\ \{(0,1),(0,-1)\}, \\ +\infty, & v \notin \text{co}\ \{(0,1),(0,-1)\}, \end{cases}$$

Example 1.

As

$$\partial f_1(x^*) = \text{co } \{(1,0), (-1,0)\},$$

$$\partial f_2(x^*) = \text{co } \{(0,1), (0,-1)\},$$

then

$$0_2 \in \partial f_1(x^*) \cap \partial f_2(x^*) = \text{co } \{(1,0), (-1,0)\} \cap \text{co } \{(0,1), (0,-1)\}.$$

Therefore $f(x^*) = 0 = f^o(0_2)$.

The point x^* is a Clarke's stationary point of the function f on $\mathbb{R}^n$. The next relation

$$\partial f_1^*(v^*) \cap \partial f_2^*(v^*) \neq \emptyset \quad \forall v^* \in \partial f_1(x^*) \cap \partial f_2(x^*)$$

is valid

Example 2.

Consider the function

$$f_1(x) = \begin{cases} x^3, & x \geq 0, \\ 0, & x < 0, \end{cases}, \quad f_2(x) = \begin{cases} 0, & x \geq 0, \\ -x^3, & x < 0, \end{cases} \quad x \in \mathbb{R}.$$

Note that the functions f_1 and f_2 are continuously differentiable. Then $f(x) = x^3$. The function f has a stationary point $x^* = 0$ and

$$f_1'(x^*) = 0, \quad f_2'(x^*) = 0.$$

Example 2.

Calculate f_1^* and f_2^*. We have

$$f_1^*(v) = \begin{cases} \dfrac{2v\sqrt{v}}{3\sqrt{3}}, & v \geq 0, \\ +\infty, & v < 0, \end{cases} \quad , \quad f_2^*(v) = \begin{cases} +\infty, & v > 0, \\ -\dfrac{2v\sqrt{|v|}}{3\sqrt{3}}, & v \leq 0, \end{cases} \quad v \in \mathbb{R}.$$

Therefore,

$$f^o(v) = \begin{cases} +\infty, & v > 0, \\ 0, & v = 0, \\ -\infty, & v < 0. \end{cases}$$

Example 2.

The function f^o is finite only in a single point $v^* = 0$. Find subdifferentials of functions f_1^* and f_2^* at the point v^*

$$\partial f_1^*(0) = (-\infty, 0] \subset \mathbb{R}, \quad \partial f_2^*(0) = [0, +\infty) \subset \mathbb{R}.$$

It is obvious that $\partial f_1^*(0) \cap \partial f_2^*(0) = 0$.
Note the fact that if we calculate the function conjugate to the function f, then $f^*(v) = +\infty \quad \forall v \in \mathbb{R}$.

1) if dom $f_2^* \not\subset$ dom f_1^*, then the function f is unbounded from below,
2) if dom $f_1^* \not\subset$ dom f_2^*, then the function f is unbounded from above.

Note that in the points not belonging to the set dom f_1^*, we face with the case $+\infty - \infty$, therefore, under minimizing the function f on $\mathbb{R}^n$, we define the function f^o on the complement of the set dom f_1^* to the whole space by the value $+\infty$.
Namely, put

$$f_-^o(v) = \begin{cases} f^o(v), & v \in \text{dom } f_1^*, \\ +\infty, & v \notin \text{dom } f_1^*. \end{cases}$$

Theorem.

Let a point $x^* \in \mathbb{R}^n$ *be a global minimizer of the function* f on $\mathbb{R}^n$, then any subgradient $v^* \in \partial f_2(x^*)$ is a global minimizer of the function f_-^o on $\mathbb{R}^n$.

Theorem.
Let a point $x^* \in \mathbb{R}^n$ be Clark's stationary point of the function f *on* $\mathbb{R}^n$. Then, if there is a global minimizer

$$v^* \in \partial f_1(x^*) \cap \partial f_2(x^*),$$

of the function f_-^o on $\mathbb{R}^n$, then the point x^* is a global minimizer of the function f *on* $\mathbb{R}^n$.

Corollary.

If there is a global minimizer $x^* \in \mathbb{R}^n$ of the function f on $\mathbb{R}^n$ then there is a point $v^* \in \mathbb{R}^n$, which is a global minimizer of the function f_-^o on $\mathbb{R}^n$. In this case the relation

$$\min_{x \in \mathbb{R}^n} f(x) = \min_{v \in \mathbb{R}^n} f_-^o(v)$$

is valid.

Corollary.

If the function f achieves at some point $x^* \in \mathbb{R}^n$ its global minimum on $\mathbb{R}^n$, then

$$\text{dom}\ \ f_2^* \subset \text{dom}\ f_1^*.$$

As it will follow from Example 1, this condition is a necessary condition.

Example 3

Consider the function

$$f(x) = f_1(x) - f_2(x) = 0.5x^2 - |x - 0.5|, \quad x \in \mathbb{R}.$$

Then the function $f(x)$ are three Clark's stationary points:

$$x_1 = -1, \quad x_2 = 0.5, \quad x_3 = 1.$$

The point x_1 is a global of f, the point x_2 is a local maximizer of f and the point x_3 is a local minimizer of f on $\mathbb{R}$ and

$$f(x_1) = -1, \quad f(x_2) = 0.125, \quad f(x_3) = 0.$$

In this problem $\mathbb{R} = \text{dom } f_1^* \not\subset \text{dom } f_2^* = [-1, 1]$.
From this statement it follows that the function f is unbounded from above.

As

$$f_1^*(v) = 0.5v^2, \quad f_2^*(v) = \begin{cases} 0.5v, & v \in [-1,1], \\ +\infty, & v \notin [-1,1], \end{cases}$$

then

$$f^o(v) = \begin{cases} 0.5v(1-v), & v \in [-1,1], \\ +\infty, & v \notin [-1,1]. \end{cases}$$

The function f^o has also three Clark's stationary points:

$$v_1 = -1,\ v_2 = 0.5,\ v_3 = 1,\ f^o(v_1) = -1,\ f^o(v_2) = 0.125,\ f^o(v_3) = 0.$$

The point v_1 is a global minimizer, the point v_2 is a local maximizer, the points v_3 is a local minimizer of f^o on $\mathbb{R}$. We have

$$v_1 = f_1'(x_1) = f_2'(x_1), \quad v_2 = f_1'(x_2) \subset \partial f_2(x_2), \quad v_3 = f_1'(x_3) = f_2'(x_3).$$

Polyhedral functions

A function

$$f(x) = \max_{i \in I}\{\langle a_i, x\rangle + b_i\}, \quad a_i \in \mathbb{R}^n, \quad b_i \in \mathbb{R}, \quad i \in I = 1, \dots, m,$$

is called a polyhedral function.

The subdifferential of a polyhedral function is a convex polyhedron, namely,

$$\partial f(x) = \text{co}\left\{\bigcup_{i \in R(x)} a_i\right\},$$

where

$$R(x) = \{i \in I \mid f_i(x) = f(x)\}, \quad f_i(x) = \langle a_i, x\rangle + b_i, \quad i \in I.$$

Note, that the subdifferential mapping $\partial f : \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^n}$ is not continuous in the Hausdorff metric.

For the function f at each point $x \in \mathbb{R}^n$ for any $\varepsilon \geq 0$ there exists the ε-subdifferential, and the ε-subdifferential mapping

$$\partial_\varepsilon f : \ \mathbb{R}^n \times (0, +\infty) \longrightarrow 2^{\mathbb{R}^n}$$

is already continuous in the Hausdorff metric.
For the polyhedral function, the formula of the ε-subdifferential at each point $x \in \mathbb{R}^n$ is

$$\partial_\varepsilon f(x) = \left\{ v = \sum_{i=1}^m \lambda_i a_i \in \mathbb{R}^n \,\middle|\, \begin{array}{c} \sum\limits_{i=1}^m \lambda_i (f(x) - \langle a_i, x\rangle - b_i) \leq \varepsilon, \\ \sum\limits_{i=1}^m \lambda_i = 1, \ \lambda_i \geq 0, \ i \in I \end{array} \right\}.$$

Thus, the ε-subdifferential of a polyhedral function f at each point $x \in \mathbb{R}^n$ is a convex polyhedron.

Remark. It is necessary to note that the points $a_i, \ i \in I$, belong to the set $\partial_\varepsilon f(x)$ for all $\varepsilon \geq f(x) - \langle a_i, x \rangle - b_i$.

Using the conjugate function, it is possible to give another definition of the ε-subdifferential of a convex closed function f at a point $x \in \mathrm{dom} f$:

$$\partial_\varepsilon f(x) = \{v \in \mathbb{R}^n \mid f(x) + f^*(v) - \langle x, v\rangle \leq \varepsilon\}.$$

The effective domain of the conjugate function f^* is the convex hull of the vectors a_i, $i \in I$, i.e.,

$$\mathrm{dom} f^* = \mathrm{co}\left\{\bigcup_{i\in I} a_i\right\}.$$

Thus, the conjugate function is finite only at points of this polyhedron. Outside of it the function f^* takes the value $+\infty$.

V.F. Demyanov introduced the notions of hypodifferentiable function and hypodifferential

A function f is called a hypodifferentable function at a point $x \in \mathbb{R}^n$ if there exists a convex compact set $df(x) \subset \mathbb{R}^{n+1}$ such that

$$f(x+\Delta) = f(x) + \max_{[a,v]\in df(x)} [a + \langle v, \Delta\rangle] + o(x,\Delta), \ a \in \mathbb{R}, \ v \in \mathbb{R}^n,$$

where

$$\frac{o(x,\alpha\Delta)}{\alpha} \longrightarrow 0 \quad \text{if} \quad \alpha \to 0 \quad \forall \Delta \in \mathbb{R}^n.$$

The set $df(x)$ is called a hypodifferential of the function f at a point $x \in \mathbb{R}^n$.

The hypodifferential of a functions f at a point $x \in \mathbb{R}^n$ is not uniquely defined. A function f is called continuously hypodifferentiable at a point $x \in \mathbb{R}^n$, if it is hypodifferentiable at x and in some neighborhood of this point there exists a continuous (in the Hausdorff metric) hypodifferentiable mapping $df(x)$. A polyhedral function is continuously hypodifferentiable in $\mathbb{R}^n$.
For example, the set

$$df(x) = \text{co}\left\{\bigcup_{i \in I}\begin{pmatrix} a_i \\ \langle a_i, x\rangle + b_i - f(x)\end{pmatrix}\right\} \subset \mathbb{R}^n \times \mathbb{R}. \tag{10}$$

can be used as a hypodifferential of the polyhedral function f.
The given hypodifferentiable mapping

$$df : \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^{n+1}}$$

is continuous in the Hausdorff metric.

The set $df(x) \subset \mathbb{R}^{n+1}$ is also a convex polyhedron contained in the half-space

$$H = \{z = (z_1 ..., z_n, z_{n+1})^T \in \mathbb{R}^n \times \mathbb{R} \mid \ z_{n+1} \leq 0\}.$$

where T denotes transposition.
For a polyhedral function f at a point $x \in \mathbb{R}^n$ we shall define the number $\varepsilon^*(x) \geq 0$ by the formula

$$\varepsilon^*(x) = \max_{i \in I}\{f(x) - f_i(x)\}. \tag{11}$$

Fix any ε, satisfying the condition $0 \le \varepsilon \le \varepsilon^*$. Put

$$d_\varepsilon f(x) = \left\{ z \in \mathbb{R}^{n+1} \middle| \ z \in df(x), z = \begin{pmatrix} v \\ t \end{pmatrix}, v \in \mathbb{R}^n, \ t \in \mathbb{R}, \ -\varepsilon \le t \le 0 \right. \tag{12}$$

The set $d_\varepsilon f(x)$ is closed and convex for any $\varepsilon: \ 0 \le \varepsilon \le \varepsilon^*$. It is not difficult to see, that

$$d_{\varepsilon_1} f(x) \subset d_{\varepsilon_2} f(x), \quad 0 \le \varepsilon_1 \le \varepsilon_2 \le \varepsilon^*.$$

Lemma.

For any $0 \leq \varepsilon \leq \varepsilon^*(x)$ *the equality*

$$\partial_\varepsilon f(x) = \left\{ v \in \mathbb{R}^n \mid \begin{pmatrix} v \\ t \end{pmatrix} \in d_\varepsilon f(x) \right\} \tag{13}$$

is valid.

Thus, the projection the set $d_\varepsilon f(x)$ onto $\mathbb{R}^n \times 0$ is the set $\partial_\varepsilon f(x) \times 0$.

Corollary.

For each $\varepsilon \geq \varepsilon^*(x)$ the equality

$$\partial_\varepsilon f(x) = \partial_{\varepsilon^*(x)} f(x) = \mathrm{co}\left\{\bigcup_{i \in I} a_i\right\} = \mathrm{dom} f^*.$$

holds.

Corollary .

If $v \notin \partial_\varepsilon f(x)$, then the point $z_t = \begin{pmatrix} v \\ t \end{pmatrix} \notin d_\varepsilon f(x)$ for every $t \in [-\varepsilon, 0]$.

Example 4

Let

$$f(x) = |x| = \max\{\ x, -x\ \},\ \ x \in \mathbb{R}.$$

Then

$$\text{dom } f^* = \text{co } \{-1, 1\} \subset \mathbb{R}.$$

Using the formula (10) calculate a hypodifferential of the function f at the point $x \in \mathbb{R}$

$$df(x) = \text{co}\left\{ \begin{pmatrix} 1 \\ x - |x| \end{pmatrix}, \begin{pmatrix} -1 \\ -x - |x| \end{pmatrix} \right\} \subset \mathbb{R}^2.$$

Example 4

If $x = 0$ then

$$df(0) = \text{co}\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\} \subset \mathbb{R}^2.$$

In this case we have from (11) $\varepsilon^*(0) = 0$. Hence,

$$\partial_\varepsilon f(0) = \partial_{\varepsilon^*(0)} f(0) = \partial f(0) \quad \forall \varepsilon \geq 0.$$

Example 4

If $x = 1$, then

$$df(1) = \text{co}\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -2 \end{pmatrix} \right\} \subset \mathbb{R}^2,$$

and $\varepsilon^*(1) = 2$. Thus,

$$\partial_\varepsilon f(1) = \text{co}\{1 - \varepsilon, 1\} \subset \mathbb{R}, \quad 0 \leq \varepsilon < \varepsilon^*(1),$$

$$\partial_\varepsilon f(1) = \partial_{\varepsilon^*(1)} f(1) = \partial f(0) = \text{dom } f^* \qquad \forall \varepsilon \geq \varepsilon^*(1).$$

Example 4.

If $x = -1$, then

$$df(-1) = \text{co}\left\{ \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\} \subset \mathbb{R}^2,$$

and $\varepsilon^*(-1) = 2$. Thus,

$$\partial_\varepsilon f(-1) = \text{co}\{-1, -1 + \varepsilon\} \subset \mathbb{R}, \quad 0 \le \varepsilon < \varepsilon^*(-1),$$

$$\partial_\varepsilon f(-1) = \partial_{\varepsilon^*(-1)} f(-1) = \partial f(0) = \text{dom } f^* \qquad \forall \varepsilon \ge \varepsilon^*(-1).$$

Example 5.

Let

$$f(x) = \max\{x+1, 2x\}, \ x \in \mathbb{R}.$$

For the given function dom $f^* = \text{co}\{1, 2\} \subset \mathbb{R}$. Using the formula (10), we have

$$df(x) = \text{co}\left\{ \begin{pmatrix} 1 \\ x+1-f(x) \end{pmatrix}, \begin{pmatrix} 2 \\ 2x-f(x) \end{pmatrix} \right\} \subset \mathbb{R}^2.$$

If $x = 1$, then

$$df(1) = \text{co}\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix} \right\} \subset \mathbb{R}^2,$$

and $\varepsilon^*(1) = 0$. Hence,

$$\partial_\varepsilon f(1) = \partial_{\varepsilon^*(1)} f(1) = \partial f(1) = \text{co}\{1, 2\} \quad \forall \varepsilon \geq 0.$$

Example 5.

If $x = 2$, then

$$df(x) = \text{co}\left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}2\\0\end{pmatrix}\right\} \subset \mathbb{R}^2,$$

and $\varepsilon^*(2) = 1$. Thus,

$$\partial_\varepsilon f(2) = \text{co}\{2-\varepsilon, 2\} \subset \mathbb{R}, \quad 0 \le \varepsilon < \varepsilon^*(2),$$

$$\partial_\varepsilon f(2) = \partial_{\varepsilon^*(2)} f(2) = \partial f(1) = \text{dom } f^* \quad \forall \varepsilon \ge \varepsilon^*(2).$$

Example 5.

If $x = 0$ then

$$df(0) = \text{co}\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix} \right\} \subset \mathbb{R}^2.$$

Then from (11) we have $\varepsilon^*(0) = 1$. Therefore

$$\partial_\varepsilon f(0) = \text{co}\{1, 1 + \varepsilon\} \subset \mathbb{R}, \quad 0 \le \varepsilon < \varepsilon^*(0),$$

$$\partial_\varepsilon f(0) = \partial_{\varepsilon^*(0)} f(0) = \partial f(1) = \text{dom } f^* \quad \forall \varepsilon \ge \varepsilon^*(0).$$

Let's denote

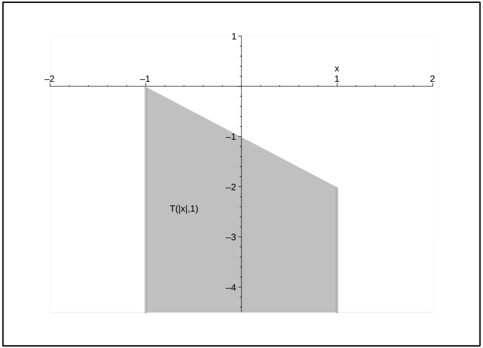
$$T(f,x) = df(x) + K, \quad T_\varepsilon(f,x) = T(f,x) \cap H(\varepsilon),$$

where

$$K = \{g \in \mathbb{R}^{n+1} \mid g = \lambda e, \; e = (\underbrace{0...0}_{n}, -1)^T, \; \lambda \geq 0\},$$

$$H(\varepsilon) = \{z = (z_1 ..., z_n, z_{n+1})^T \in \mathbb{R}^{n+1} \mid z_{n+1} = -\varepsilon\}.$$

A geometrical interpretation of the ε-subdifferential of polyhe



$$H(\varepsilon)$$

Lemma

For any fixed $\varepsilon \geq 0$ at each point $x \in \mathbb{R}^n$ the equality

$$\partial_\varepsilon f(x) = \left\{ v \in \mathbb{R}^n \mid \begin{pmatrix} v \\ t \end{pmatrix} \in T_\varepsilon(f, x) \right\} \quad (14)$$

holds.

Let's consider the optimization problem: find

$$\inf_{x \in \mathbb{R}^n} f(x).$$

Let f_1 and f_2 be the polyhedral functions defined on $\mathbb{R}^n$, i.e.

$$f_1(x) = \max_{i \in I} f_{1i}(x), \ f_{1i} = \{\langle a_i, x\rangle + b_i\}, \ I = \{1, \dots, m\},$$

$$f_2(x) = \max_{j \in J} f_{2j}(x), \ f_{2j}(x) = \{\langle c_j, x\rangle + d_j\}, \ J = \{1, \dots, p\},$$

Where $a_i, \ c_j \in \mathbb{R}^n, \ b_i, \ d_j \in \mathbb{R}, \ i \in I, \ j \in J.$

Consider the function $f(x) = f_1(x) - f_2(x)$. Then

$$f(x) = \max_{i \in I}\{\langle a_i, x\rangle + b_i\} - \max_{j \in J}\{\langle c_j, x\rangle + d_j\}.$$

Consider some optimization properties of the function f.
Reduce conditions of the unboundedness of the function f on $\mathbb{R}^n$.

Theorem.

For the function f to be unbounded from below in $\mathbb{R}^n$, it is necessary and sufficient that there exist $j^* \in J$ and a vector c_{j^*}, such that the condition

$$c_{j^*} \notin \operatorname{co}\left\{\bigcup_{i \in I} a_i\right\} \tag{15}$$

holds.

Corollary

For the function f to be unbounded from below in $\mathbb{R}^n$, it is necessary and sufficient that the condition

$$\text{dom } f_2^* \not\subset \text{dom } f_1^*$$

hold.

Corollary

For the function f to be bounded from below in $\mathbb{R}^n$, it is necessary and sufficient that the inclusion

$$\text{dom } f_2^* \subset \text{dom } f_1^* \tag{16}$$

be valid.

Theorem

For the function f to be unbounded from above in $\mathbb{R}^n$, it is necessary and sufficient that there exist $i^* \in I$ and a vector a_{i^*}, such that the condition

$$a_{i^*} \notin \operatorname{co}\left\{\bigcup_{j \in J} c_j\right\}$$

be satisfied.

Corollary

For the function f to be unbounded from above in $\mathbb{R}^n$, it is necessary and sufficient that the condition

$$\text{dom } f_1^* \not\subset \text{dom } f_2^*$$

hold.

Corollary

For the function f to be bounded from above in $\mathbb{R}^n$, it is necessary and sufficient that the inclusion

$$\text{dom } f_1^* \subset \text{dom } f_2^*.$$

be valid.

Let the function f be bounded from below in $\mathbb{R}^n$, i.e., the condition (16) takes place.

Theorem

For the point $x^* \in \mathbb{R}^n$ be a global minimizer of the function f on $\mathbb{R}^n$, it is necessary and sufficient, that the condition

$$df_1(x^*) \bigcap \mathrm{co} \left\{ \begin{pmatrix} c_j \\ f_{2j}(x^*) - f_2(x^*) \end{pmatrix}, \begin{pmatrix} c_j \\ 0 \end{pmatrix} \right\} \neq \emptyset \quad \forall j \in J, \qquad (17)$$

hold.

Corollary

The condition (17) *is equivalent to the following condition*

$$0_{n+1} \in \left[df_1(x^*) - \operatorname{co}\left\{ \begin{pmatrix} c_{j^*} \\ f_{2j^*}(x^*) - f_2(x^*) \end{pmatrix}, \begin{pmatrix} c_{j^*} \\ 0 \end{pmatrix} \right\} \right] \quad \forall j \in J.$$

Corollary

(The sufficient condition for a global minimum of the function f on $\mathbb{R}^n$.) *If at a point $x^* \in \mathbb{R}^n$ the inclusion*

$$df_2(x^*) \subset df_1(x^*)$$

holds then the point x^* is a global minimizer of the function f on $\mathbb{R}^n$.

Example 6.

Consider the function

$$f(x) = f_1(x) - f_2(x),$$

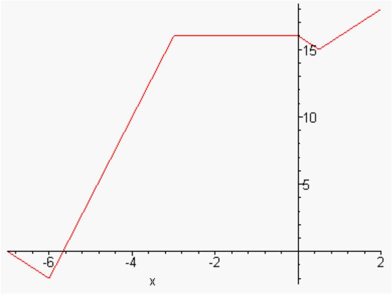
where

$$f_1(x) = \max\{|6x+23|, |2x+25|\}, \qquad f_2(x) = \max\{|4x+9|, |2x+9|\}$$

We have

$$f(x) = \begin{cases} -14-2x, & \text{if } -\infty < x \le -6, \\ 34-6x, & \text{if } -6 < x \le -3, \\ 16, & \text{if } -3 < x \le 0, \\ 16-2x, & \text{if } 0 < x \le \frac{1}{2}, \\ 2x+14, & \text{if } \frac{1}{2} < x \le +\infty. \end{cases}$$

On fig. 1 the function f is represented.



It is easy to see, that

$$\mathrm{dom} f_1^* = \mathrm{co}\{-6, 6\}, \quad \mathrm{dom} f_2^* = \mathrm{co}\{-4, 4\}.$$

Thus, this function is bounded from below ($\mathrm{dom} f_2^* \subset \mathrm{dom} f_1^*$) and unbounded from above. The point $x^* = -6$ is a global minimizer of the function f on $\mathbb{R}$. At this point $f(-6) = -2$ and

$$\partial f_1(-6) = \mathrm{co}\{-6, 2\}, \; df_1(-6) = \mathrm{co}\left\{\begin{pmatrix} 6 \\ -26 \end{pmatrix}, \begin{pmatrix} -2 \\ -26 \end{pmatrix}, \begin{pmatrix} -6 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}\right\},$$

$$\partial f_2(-6) = -4, \quad df_2(-6) = \mathrm{co}\left\{\begin{pmatrix} 4 \\ -30 \end{pmatrix}, \begin{pmatrix} 2 \\ -18 \end{pmatrix}, \begin{pmatrix} -4 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 12 \end{pmatrix}\right\}.$$

It is obvious, that the condition (4) holds. See fig.2.

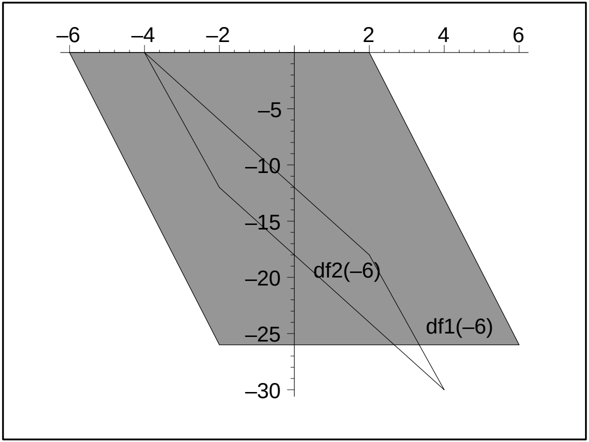


Fig. 2.

Any point from the interval $(-3, 0)$ is a stationary point of the function f. The functions f_1 and f_2 are differentiable on this interval and $f_1'(x) = 2,\ f_2'(x) = 2$ for any $x \in (-3, 0)$.
Consider the point $x_1 = -2$. We have

$$df_1(-2) = \text{co}\left\{ \begin{pmatrix} 6 \\ -10 \end{pmatrix}, \begin{pmatrix} -2 \\ -42 \end{pmatrix}, \begin{pmatrix} -6 \\ -32 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix} \right\},$$

$$df_2(-2) = \text{co}\left\{ \begin{pmatrix} 4 \\ -4 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} -4 \\ -6 \end{pmatrix}, \begin{pmatrix} -2 \\ -10 \end{pmatrix} \right\}.$$

The condition (4) holds, but the condition (17) does not. See fig 3.

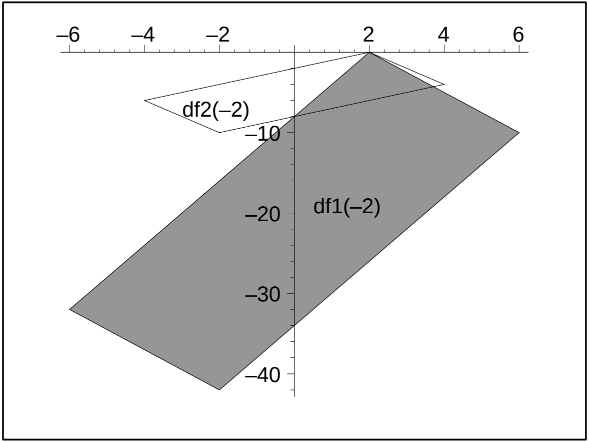


Fig. 3.

Take the point $x_2 = \frac{1}{2}$. This is a strict local minimizer of the function f. Then $f\left(\frac{1}{2}\right) = 15$ and

$$\partial f_1\left(\frac{1}{2}\right) = \text{co}\,\{2,6\}\,,\; df_1(-6) = \text{co}\left\{\begin{pmatrix}6\\0\end{pmatrix},\begin{pmatrix}-2\\-52\end{pmatrix},\begin{pmatrix}-6\\-52\end{pmatrix},\begin{pmatrix}2\\0\end{pmatrix}\right\},$$

$$\partial f_2\left(\frac{1}{2}\right) = 4, \quad df_2(-6) = \text{co}\left\{\begin{pmatrix}4\\0\end{pmatrix},\begin{pmatrix}2\\-1\end{pmatrix},\begin{pmatrix}-4\\-22\end{pmatrix},\begin{pmatrix}-2\\-21\end{pmatrix}\right\}.$$

Observe, that $\partial f_2\left(\frac{1}{2}\right) \subset \text{int}\ \partial f_2\left(\frac{1}{2}\right)$, i.e., the sufficient condition for a strict local minimum is satisfied. The condition (17) does not hold.

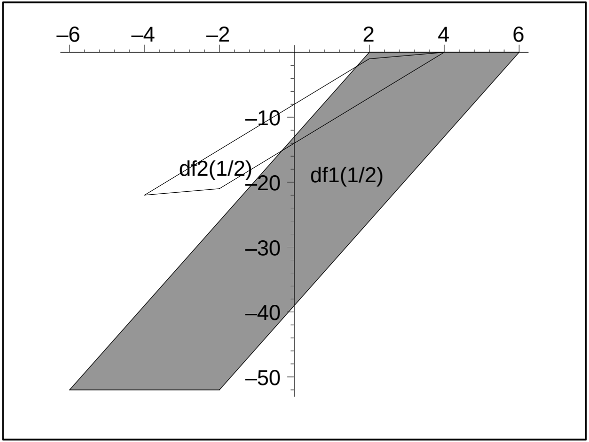


Fig. 4.